The Thesis committee for Inh Jee

Certifies that this is the approved version of the following thesis:

Measuring Angular Diameter Distances in the Universe

by Baryon Acoustic Oscillation and Strong Gravitational Lensing

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor: _____

Karl Gebhardt

_____

Steven Finkelstein

Measuring Angular Diameter Distances in the Universe by Baryon Acoustic Oscillation and Strong Gravitational Lensing

by

Inh Jee, B.S.

Thesis

Presented to the Faculty of the Graduate School

of The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

August 2013

Acknowledgments

We would like to acknowledge Eiichiro Komatsu for his thorough advising on the author's academic work as an external supervisor, Donghui Jeong for his help on creating the code for generating the HETDEX mock geometry and log-normal realizations, as well as academic advices, Chi-Ting Chiang for providing the 1000 realizations with his sparse-sampling strategy he has optimized and various discussions, and Edward L. Robinson for advises on the statistical aspect of this study. Funding for this work has provided in parts by the University of Texas College of Natural Sciences and the Texas Cosmology Center through grants to Eiichiro Komatsu and Karl Gebhardt, and in part by the National Science Foundation award AST-0926815 for Hobby-Eberly Telescope Dark Energy Experiment.

Inh Jee

The University of Texas at Austin

August 2013

Measuring Angular Diameter Distances in the Universe

by Baryon Acoustic Oscillation and Strong Gravitational Lensing

by

Inh Jee, M.A.
The University of Texas at Austin, 2013
SUPERVISOR: Karl Gebhardt

We discuss two ways of measuring angular diameter distances in the Universe: (i) Baryon Acoustic Oscillation (BAO) , and (ii) strong gravitational lensing.

For (i), we study the effects of survey geometry and selection functions on the 2-point correlation function of Lyman-α emitters in $1.9 < z < 3.5$ for Hobby-Eberly Telescope Dark Energy Experiment (HETDEX). We develop a method to extract the BAO scale (hence a volume-averaged angular diameter distance D_V, which is a combination of the angular diameter distance and the Hubble expansion rate, i.e., $[cz(1+z)^2 D_A^2 H^{-1}]^{1/3}$) from a spherically averaged 1-d correlation function. We quantify the statistical errors on such measurements. By using log-normal realizations of the HETDEX dataset, we show that we can determine D_V from HETDEX at 2% accuracy using the 2-point correlation function. This study is complementary to the on-going effort to characterize the power spectrum using HETDEX.

For (ii), a previous study (Paraficz & Hjorth, 2009) looked at the case of a spherical lens following a singular isothermal distribution of matter and an isotropic velocity distribution, and found that combining measurements of the Einstein ring radius with the time delay of a strong lens system directly leads to a measurement of the angular diameter distance, D_A. Since this is a very new method, it requires more careful investigations of various real-world effects such as a realistic matter

density profile, an anisotropic velocity distribution, and external convergence. In more realistic lens configurations we find that the velocity dispersion is the dominant source of the uncertainty ; in order for this method to achieve competitive precision on measurements of D_A, we need to constrain the velocity dispersion, down to the percent level. On the other hand, external convergence and velocity dispersion anisotropy have negligible effect on our result. However, we also claim that the dominant source of the uncertainty depends largely on the image configuration of the system, which leads us to the conclusion that studying the angular dependence of the lens mass distribution is a necessary component.

Contents

Acknowledgments **iv**

Abstract **v**

Chapter 1 Introduction **1**

Chapter 2 Constraining the angular diameter distance from BAO **5**

2.1 BAO 5

2.2 HETDEX 7

2.3 2-point correlation function 12

2.4 Simulation 16

2.5 Results 17

2.6 2-dimensional 2-point correlation function and the Alcock-Paczyński (AP) Test 24

Chapter 3 Strong Lensing as a Standard Ruler **29**

3.1 Physical motivations 29

3.2 Simple analysis using a singular isothermal sphere 30

3.3 More realistic analysis 33

3.4 Error formula and implications for B1608+686 and RJX1131-1231 . 50

3.5 Conclusion 60

Bibliography **62**

Chapter 1

Introduction

After the discovery of the accelerated expansion of the universe, the distance-redshift relation became an important tool for constraining cosmological model, and understanding the nature of dark energy. On cosmological scale there are two known methods to measure the distances. One is by using objects with a known brightness, called the *standard candle method*. Type Ia supernovae are a good example of standard candle which allowed the first detection of the accelerated expansion of the universe. The other method is by using an object with known size, called the *standard ruler method*. In this thesis, we study two different methods as a standard ruler.

For the first part of this thesis, we study Baryon Acoustic Oscillation (BAO). BAO is widely used as a standard ruler in measuring the angular diameter distances and Hubble expansion rates at different redshifts. The BAO originates in photon-

baryon interaction in the early epoch of the universe before decoupling happened at $z = 1090$, and its characteristic scale known as the *sound horizon scale*, has been precisely measured by Cosmic Microwave Background(CMB) probes, such as the *Wilkinson Microwave Anisotropy Probe* (*WMAP*). However, since BAO is a generic result of interactions between photons and baryons, the BAO feature exists not only in the CMB but also in the matter distribution, and since galaxies are collapsed objects growing inside over dense regions of the matter distribution, we can use them as tracers of matter distribution by assuming some appropriate galaxy bias. For example, Sloan Digital Sky Survey (SDSS) team was able to detect the BAO feature in the power spectrum and 2-point correlation function of their 46,000 luminous red galaxies (LRGs) samples at $0.16 < z < 0.47$ (Eisenstein et al., 2005), and thus could constrain the distance to the corresponding redshift.

While a galaxy survey is a good way to measure D_A, it requires substantial resources and many years of dedicated observations. Is there an alternative way to measure D_A? For the second part of this thesis, we study such a possibility using strong gravitational lensing. It has long been known that systems involving strong gravitational lens systems can be used to determine cosmological parameters via their sensitivity to a distance-like quantity, which is the *ratio* of distances in a strong lens system (Fukugita et al., 1990; Fukugita & Turner, 1991; Futamase & Yoshida, 2001; Yamamoto & Futamase, 2001; Yamamoto et al., 2001; Ohyama et al.,

2002; Dobke et al., 2009; Suyu et al., 2009). These different distance ratios can be calculated using observables such as the frequency of lens systems (Fukugita et al., 1990; Fukugita & Turner, 1991); Einstein ring radius (Futamase & Yoshida, 2001; Yamamoto & Futamase, 2001; Yamamoto et al., 2001; Ohyama et al., 2002; Dobke et al., 2009); and lensing time delay Δt(Oguri, 2007; Coe & Moustakas, 2009; Suyu et al., 2009). However, Paraficz & Hjorth (2009) recently observed that the actual angular diameter distance to the lens could be determined by measuring the time delay between two spatially separated images of a single source via strong lensing, and measuring the velocity dispersion of the lens galaxy.

Their idea has the potential to become a new standard ruler for measuring cosmological distances. In this study, we formulate a way to determine D_A from strong lens systems. Specifically, we combine two lensing observables: separation of images (e.g., Einstein ring radius for a spherical system) and time delay under a specific mass model of lens galaxy. Also, we investigate the lens galaxy parameters to find the dominant source of the uncertainty in D_A derived from our model. We find that improving the velocity dispersion measurement of lens galaxy will allow us to constrain D_A to a few percent accuracy.

The structure of this paper is as follows. In Section 2, we explain what BAO is in detail, and show how we use it to constrain D_V using the 2-point correlation function. More specifically, we calculate the 2-point correlation function of

simulated galaxy data consistent with the HETDEX survey geometry to show the reliability of the method. In Section 3, we briefly summarize the previous study of Paraficz & Hjorth (2009) that used strong gravitational lensing system as a standard ruler, and then extend their model by including several realistic parameters : i) a power-law density profile, ii) an anisotropic velocity dispersion, and iii) external convergence. We also study the effects of uncertainties in observables and errors on lens model parameters on angular diameter distances. Then we apply the analysis to two well-studied time-delayed strong gravitational lensing systems, B1608+656 and RXJ1131-1231, to numerically check the reliability of our methods. We conclude in Section 4.

Chapter 2

Constraining the angular diameter distance from BAO

2.1 BAO

BAO is the result of photon-baryon coupling at redshifts higher than $z = 1090$, when baryons and photons decouple. During that epoch, baryons and photons are tightly coupled to each other via Thompson scattering, and behave like a single fluid. This photon-baryon fluid, due to the large pressure provided by the photons, cannot collapse under gravity but instead oscillates, resulting in the so-called BAO. The characteristic scale of this oscillation is determined by the sound horizon r_s at the decoupling epoch, given by

$$r_s \equiv \int_0^{t_{\rm dec}} c_s(t) \frac{dt}{a(t)}, \tag{2.1}$$

where

$$c_s(t) \equiv \frac{1}{\sqrt{3\left[1+\frac{3\rho_b(t)}{4\rho_\gamma(t)}\right]}} = \frac{1}{\sqrt{3\left[1+\frac{3\Omega_b}{4\Omega_\gamma}a(t)\right]}} = \frac{1}{\sqrt{3\left[1+0.67\left(\frac{\Omega_b h^2}{0.023}\right)\left(\frac{1090}{1+z}\right)\right]}}, \tag{2.2}$$

is the speed of sound of the baryon-photon fluid. Observe that the sound horizon scale is the comovig distance that a perturbation of the photon-baryon fluid could propagate through until the decoupling. Using the *WMAP* 7-year parameters, one finds $r_s = 153.2 \pm 1.7$ Mpc.

This BAO sound horizon scale is a characteristic scale observed in both the CMB and galaxies distribution. It can be used to derive the angular diameter distance D_A and Hubble expansion rate H at a desired redshift, from the observed size of BAO features from the angular galaxy distribution on the sky and the galaxy distribution along the line of sight, respectively. Specifically, we have two relations:

$$\delta\theta = \frac{1}{1+z}\frac{r_s}{D_A(z)}, \tag{2.3}$$

where $\delta\theta$ is the measured angular separation of the BAO feature in the 2-point correlation function of the galaxy distribution on the sky, and

$$c\delta z = H(z) r_s, \tag{2.4}$$

where δz is the measured redshift separation of the BAO feature in the 2-point

correlation function along the line of sight.

One can then convert these measured D_A and H into cosmological parameters (most notably, giving information about the nature of dark energy) via

$$D_A = \frac{1}{1+z}\int_0^z \frac{dz}{H(z)} \tag{2.5}$$

in a flat universe, and

$$H(z) = H_0\left[\Omega_m(1+z)^3 + \Omega_{de}(1+z)^{3(1+w)}\right]^{1/2}, \tag{2.6}$$

when we assume the equation of state for dark energy w is constant.

2.2 HETDEX

HETDEX is one of the most cost-efficient large-scale galaxy surveys to date. With 75 Visible Integral-Field Replicable Unit Spectrographs (VIRUS) attached to the 10-m Hobby-Eberly Telescope (HET) at McDonald Observatory, HETDEX is expected to detect 0.8 million Lyman alpha emitting galaxies (LAEs) at $1.9 < z < 3.5$. The principle goal of the survey is to constrain the dark energy equation of state by calculating the angular diameter distance and Hubble parameter at different redshifts from the BAO (sound horizon) scale in the 2-point correlation function and power spectrum of LAEs. However, since the sound horizon scale is 150 Mpc, we

need a wide field observation (i.e. at least a few times larger than the scale of interest), to correctly obtain the signal. Since fully observing an appropriately large area would be too demanding in terms of resources, HETDEX will instead sparsely sample the sub-domain of the sky which observation will cover (hereafter the *survey footprint*), while minimizing the statistical uncertainty in the power spectrum and 2-point correlation function. For this purpose, the HETDEX team has developed a distinct survey strategy, and also studied the impact of sparse sampling on the measured power spectrum (Chiang et al., 2013). We will introduce the survey geometry in detail in the following sections.

2.2.1 Survey footprint

The HETDEX footprint is centered at $(RA, DEC) = (13^h, 53^\circ)$ and covers 41.7×7.7 deg^2 on the sky. If we covered the entire redshift range stated in the previous section, this would correspond to a comoving volume of 10 Gpc^3. In simulations, however, we only sample the lower redshift portion ($1.9 < z < 2.5$) of the spring field of the survey in order to avoid galaxy bias evolution caused by such a large redshift range, which would distort our measurements. In comoving units, the side lengths for the simulation box are $L_x = 3021\ h^{-1}\mathrm{Mpc}$, $L_y = 540.6\ h^{-1}\mathrm{Mpc}$, and $L_z = 836.1\ h^{-1}\mathrm{Mpc}$, and the resulting volume is $1.365\ h^{-3}\mathrm{Gpc}^3$. This is about one-third of the total proposed HETDEX survey volume.

We chose the input power spectrum to be non-linear power spectrum at $z = 2.2$, based on third-order perturbation theory with non-linear bias (Jeong & Komatsu, 2006, 2009). Here, the non-linear bias is defined as

$$\delta_g(\vec{x}) = b_1\delta_m(\vec{x}) + \frac{b_2}{2}[\delta_m(\vec{x})^2 - \sigma^2], \tag{2.7}$$

where δ_g is the density contrast of the galaxy distribution, δ_m is the density contrast of the underlying matter distribution, and $\sigma^2 =< |\delta_m(\vec{x})|^2 >$. For the bias parameters, we use $b_1 = 2.2$, $b_2 = 0.671$, and $P_0 = 72.13\ h^{-3}\mathrm{Mpc}^3$.

2.2.2 Mask

Since HETDEX is a blind survey, we need spectroscopic data from all the targets to properly identify Lyman alpha emitters (hereafter LAEs) in the desired redshift range. This is possible because of the so-called Integral Field Unit(IFU) spectrographs, which takes spectra of all the objects inside the field-of-view. HETDEX will use Visible Integral-Field Replicable Unit Spectrographs (VIRUS ; Hill et al. (2010)). VIRUS have 224 fibers to feed each spectrograph, two spectrographs build one unit, and in total 75 IFUs are planned to be built and outfit the Hobby-Eberly Telescope (HET) in the configuration as shown in figure 2.1. The figure shows that we will observe only part of the area inside the telescope's field-of-view: only galaxies lying inside the field-of-view of IFUs will be used in measuring the 2-point

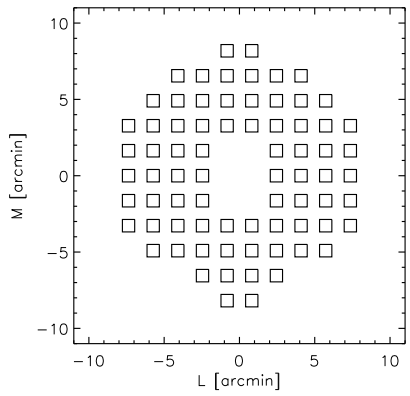

Figure 2.1 IFU locations on the focal plane of the HETDEX instrument. Each square corresponds to a single IFU, which feeds two spectrographs. The hole at the center exists to accommodate another instrument. We have 74 IFUs in this configuration. We refer to a reading taken by all the instruments while pointed at one location in the sky as a "*shot*".

correlation function.

Due to this complicated geometry, one should perform mock analyses on different survey geometry and survey strategies, to make sure that we can reconstruct the underlying 2-point correlation function by analyzing the sparsely sampled observations. For this purpose, we first generate mock galaxy distributions using the so-called "log-normal realization " (see section 2.4 for more details) method, which we will hereafter call "*no-selection*". Those act as a control sample for comparison. From no-selection galaxies, we select galaxies based on different criteria mimicking the actual survey geometry. We apply window functions to this no-selection dataset to mimic the survey geometry effect on our calculation.

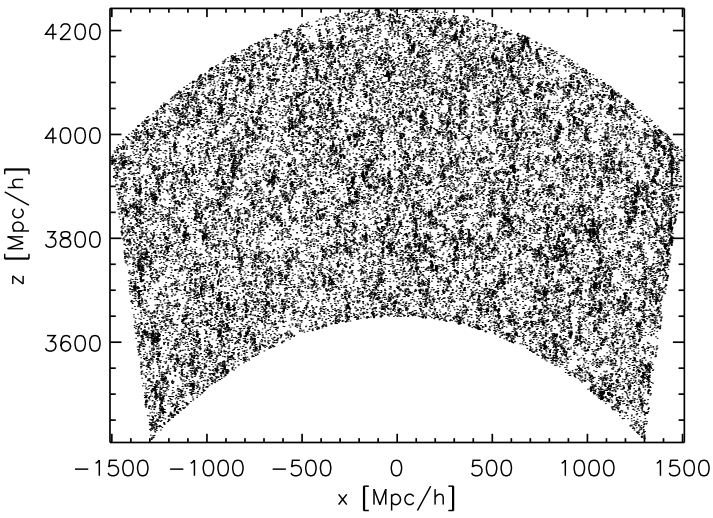

Figure 2.2 HETDEX footprint galaxies (geometric selection) in comoving coordinate, in xz plane. Due to the large survey area, the line-of-sight direction is parallel to the z-direction only at the center of the survey volume.

Our first selection criterion is "*geometric selection*", which selects all the galaxies lying inside the survey footprint. Since the survey observes the large area on the sky with constant RA and DEC, the detected volume in comoving coordinate is curved; this is especially a big problem for the power spectrum because it involves a Fourier transform. Studying geometric selection will allow us to know how non-cubical survey volume changes the 2-point correlation function.

For the second selection criterion, we choose galaxies that lie inside the IFUs. We call this "*shot+mask selection*". The number of galaxies in shot + mask selection drops to 10% level of no-selection, causing larger Poisson noise. However, unlike

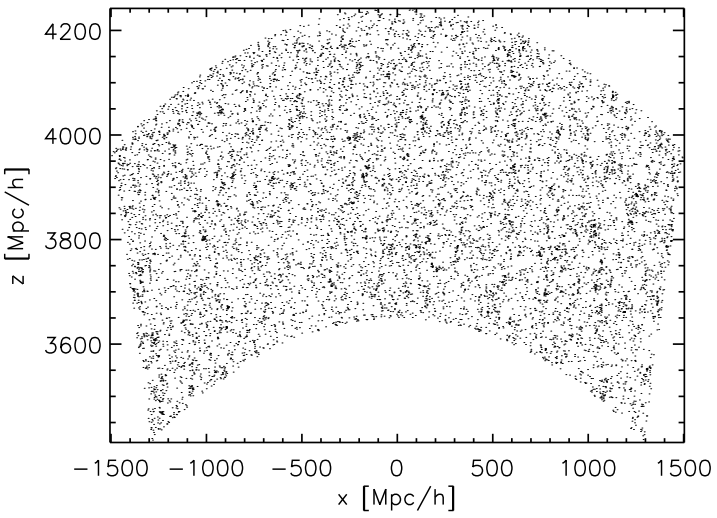

Figure 2.3 HETDEX shot + mask selection galaxies in comoving coordinate, in xz plane.

power spectrum calculations in Fourier space, the window function does not get convolved in real space. Our expectation is that the window function can change the uncertainty level of the 2-point correlation function due to the smaller number of observed galaxies but that the peak position of the BAO will not move. We show the reasoning behind this statement in the next section.

2.3 2-point correlation function

In this work, we are focused on measuring the real space (in order to measure BAO) 2-point correlation. The 2-point correlation function, usually referred to as $\xi(\vec{r})$, is defined as the volume average of the product of the density contrast $\delta(\vec{x})$ at 2 points

separated by $\vec{r} = \vec{x_1} - \vec{x_2}$,

$$\langle \delta(\vec{x_1})\ \delta(\vec{x_2}) \rangle = \xi(\vec{r}) + \frac{1}{\bar{n}} \delta_D^{(3)}(\vec{r}), \tag{2.8}$$

where the last term corresponds to the coincidence limit (i.e. when $\vec{x_1} = \vec{x_2}$). It has the following physical interpretation. The density constrast $\delta(\vec{x})$ is

$$\delta(\vec{x}) = \frac{n(\vec{x}) - \bar{n}}{\bar{n}}, \tag{2.9}$$

where $n(\vec{x})$ is the number density of galaxies at $\vec{x}$. Thus

$$\begin{aligned} \langle \delta(\vec{x_1})\ \delta(\vec{x_2}) \rangle =& \frac{\langle n(\vec{x_1})\ n(\vec{x_2}) \rangle - \bar{n}\langle n(\vec{x_1}) \rangle - \bar{n}\langle n(\vec{x_2}) \rangle + \bar{n}^2}{\bar{n}^2} \\ =& \frac{\langle n(\vec{x_1})\ n(\vec{x_2}) \rangle - \bar{n}^2}{\bar{n}^2}, \end{aligned} \tag{2.10}$$

where $\langle n(\vec{x_1})\ n(\vec{x_2}) \rangle$ is the volume average of the number of pairs of galaxies per volume, i.e. the probability to find a galaxy pair at $\vec{x_1}$ and $\vec{x_2}$. From this we can say that

$$\frac{d\text{Prob}}{d^3\vec{x_1} d^3\vec{x_2}} = \bar{n}\delta_D^{(3)}(\vec{r}) + \bar{n}^2[1 + \xi(\vec{r})]. \tag{2.11}$$

If we ignore the coincident limit, then from the equation above, $\xi(\vec{r})$ can be interpreted as the local probability of finding a galaxy pair at a separation $\vec{r}$ from the average probability.

Since BAO imprints an enhanced correlation in the matter distribution at

the sound horizon scale, we can use the 2-point correlation function as a statistical tool to detect BAO. More specifically, we expect a bump at the sound horizon scale in the 2-point correlation function.

In this study we will work with $\xi(r)$, which is the 1-dimensional 2-point correlation function, i.e. under the assumption that the galaxy distribution is spherically symmetric. For this purpose, we use two observed quantities, specifically $\delta\theta$ and δz (see equations (2.3) and (2.4)), so that what we measure from the 2-point correlation function is not D_A itself but the volume-averaged distance D_V which is defined as

$$D_V \equiv [\ cz(1+z)^2 D_A^2 H^{-1}]^{1/3} \tag{2.12}$$

(Anderson et al., 2012).

Also, since the galaxy distribution is discrete, we should choose an appropriate estimator to calculate ξ. Landy & Szalay proposed an estimator (hereafter the *LS estimator*) that is defined as

$$\xi(r) = \frac{DD - 2DR + RR}{RR} \tag{2.13}$$

(Landy & Szalay, 1993). Here, DD, DR and RR are the number of galaxy-galaxy, galaxy-random, and random-random pairs, respectively, with proper normalization. We normalize each pair count by the total possible number of pairs between the

particles in the data. Thus, for auto correlation (i.e. DD and DR), we normalize by $N(N-1)/2$ where N is the number of particles ; for cross correlation (i.e. DR), we normalize by $N_D \times N_R$ where N_D is the number of galaxies and N_R is the number of random particles. One advantage of the LS estimator is that it is an unbiased estimator for a 2-point correlation function with small variance (Landy & Szalay, 1993). Another advantage is that the survey geometry is automatically taken into account when we divide by RR. Again, the 2-point correlation function is a measure of the probability to find excess pairs of galaxies at a given separation compare to the uniform distribution. Thus, by dividing the number of galaxy pairs by the number of random pairs calculated from the same geometry, we can get rid of the selection effect such that we get the underlying 2-point correlation function.

Measuring the 2-point correlation function, $\xi(r)$ from galaxy survey data is a standard technique in cosmology, and thus there are no conceptual difficulties in measuring $\xi(r)$. In practice, however, computation of $\xi(r)$ can be quite time consuming. The time it takes to compute $\xi(r)$ from direct pair counting scales as N_g^2 (where N_g is the number of galaxies in a given survey volume) and quickly becomes very large.

We have developed a code based on the so-called *KD-tree* algorithm to efficiently compute $\xi(r)$ from a large data set such as HETDEX. This code is based upon an earlier version written (in python) by a former undergraduate student Nico-

las Canac. We debugged his code, converted it into C for better performance and memory management, parallelized it in OpenMP. As a result, the efficiency of pair counting has increased significantly : For a distribution of 0.5-million galaxies $\times$ 3-million random particles , the code takes $\sim$ 4 minutes, and for a distribution of 4-million galaxies $\times$ 6-million random particles, it takes $\sim$ 25 minutes, both cases with 16 cores. Thanks to fast pair-counting code and log-normal realizations, we can now create a large number of realizations to get statistically meaningful results, and thus estimate the 2-point correlation function in a reasonable amount of time.

2.4 Simulation

In order to study the statistical and systematic errors of the measured 2-point correlation function, we need many realizations to obtain a smooth enough covariance matrix to give us an accurate *dilation parameter*, which lets us compare our 2-point correlation function calculated from different realizations to the input 2-point correlation function. We will discuss this in more detail in the following section.

While it is expensive to run cosmological simulations many times, we can much more quickly create approximate mock catalogues using a Poisson sampling of a given density field. Specifically, we use a method called *log-normal realization*. From the input power spectrum, we first generate a density field with a given probability density distribution function (taken here to be a log-normal distribu-

tion, which guarantees positive density at all points) and then paint the galaxies on the density field by the acceptance-rejection algorithm. In this way, the power spectrum of the resulting discretely distributed points will follow that of the input power spectrum. Donghui Jeong wrote a code for generating mock catalogues, and showed that the power spectrum of a log-normal realization agrees well with that of an input power spectrum, and with small discrepancies at small scales. Chi-Ting Chiang used this code to provide us with the position data for 1000 realizations.

2.5 Results

2.5.1 Effects of survey geometry

Figure 2.4 shows the excellent agreement between the input 2-point correlation function and the measured one with no-selection. Conventionally, we multiply the 2-point correlation function by $r^{2.5}$ to highlight the BAO peak. Keeping in mind that the signal is artificially amplified toward large separation, the measured 2-point correlation function matches well with the input power spectrum.

In Figure 2.5, we show $\xi(r)$ measured from 100 log-normal realizations we created from an input power spectrum at $z = 2.2$, with 3 different selection criteria applied, as well as the input power spectrum converted into the 2-point correlation function for a reference (For this purpose, we used the *2-point Correlation Function from 3rd-order Perturbation Theory* routine in the Cosmology Routine Li-

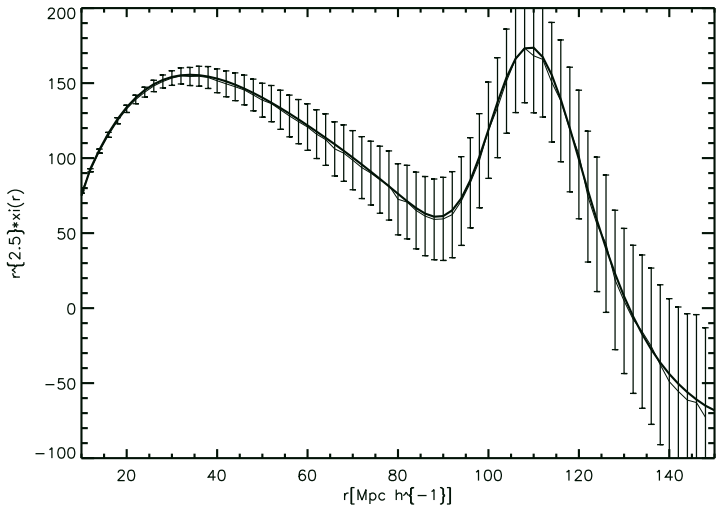

Figure 2.4 2-point correlation function $\xi(r)$ times $r^{2.5}$ as a function of r. The exponent 2.5 is conventional, to make the BAO bump pronounced. The thin solid line shows $\xi(r)$ measured from the 1000 lognormal realizations at $z = 2.2$, while the thick solid line shows the theoretical expectation, which is calculated from the input power spectrum. The errorbars represent scatters among each realizations, rather than the errorbar of the mean of 1000 realizations. This is averaged over 1000 realizations, and thus the actual data would be noisier.

brary(CRL) provided by Eiichiro Komatsu). For no-selection, there are 4-million galaxies and 3-million random particles; for geometric selection, there are 2.6-million galaxies and 3-million random particles; and for shot + mask selection there are 0.5-million galaxies and 3-million random particles.

Figure 2.5 shows that the measured 2-point correlation function has the same shape and peak position regardless of the selection choice, aside from the inevitable noise increase, i.e. Poisson error from the decreased number of galaxies in the more restrictive selections.

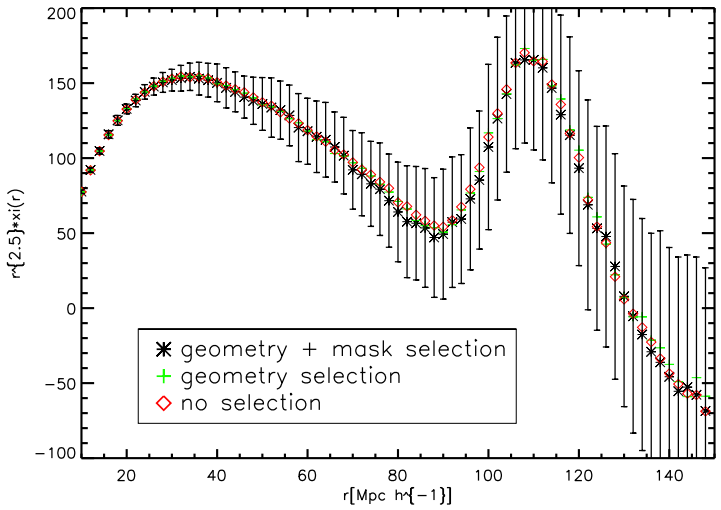

Figure 2.5 Measured 2-point correlation function, $\xi(r)$, times $r^{2.5}$ as a function of r. Red diamonds indicate no-selection, green crosses indicate geometric selection, and black stars indicate geometry + mask selection. We calculated the error bars from geometry + mask selection. This is averaged over 100 realizations.

Hereafter, we consider only shot + mask selection as that is the closest selection to the actual survey window function in terms of geometry, which we are interested in.

2.5.2 Dilation parameter

In figure 2.5, it appears that the peak position of the BAO bump averaged over 100 realizations has not shifted with respect to the input one. To quantitatively validate this statement, we define a dilation parameter α:

$$D_V/r_s = \alpha \ (D_V/r_s)_{\rm fid} \tag{2.14}$$

where $(D_V/r_s)_{\rm fid}$ on the right hand side indicates the fiducial value of D_V/r_s as calculated from the input power spectrum and WMAP 7-year cosmology while D_V/r_s on the left hand side is measured from realizations. α parametrizes the displacement of the BAO peak position from its value in the input data. We would like α to be close to unity, to validate that our calculation is not biased and that we can use the 2-point correlation function as an unbiased estimator in measuring the BAO peak. We determine α by choosing the α value that minimizes the χ^2

$$\chi^2 = [\xi^{\rm fit}(\alpha r_i) - \xi(r_i)]^T (C^{-1})_{ij} [\xi^{\rm fit}(\alpha r_j) - \xi(r_j)], \tag{2.15}$$

where C is the covariance matrix calculated from 1000 realization, $\xi^{\rm fit}$ is

$$\xi^{\rm fit}(\alpha r) = B^2 \xi_{\rm model}(\alpha r) + \frac{a_1}{r^2} + \frac{a_2}{r} + a_3 \tag{2.16}$$

and $\xi_{\rm model}$ is the reference 2-point correlation function calculated from the input power spectrum. Here, B is the linear bias, and by adding a_1, a_2 and a_3, we give extra degree of freedom to marginalize the broadband features that are not modeled in our model 2-point correlation function (Anderson et al., 2012). We explain more about the covariance matrix and the non-linear fitting method we used in the following two sections.

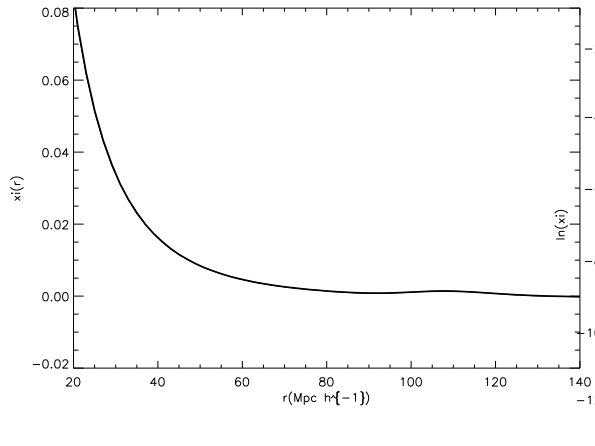

Figure 2.6 The 2-point correlation function, without multiplying by $r^{2.5}$. The thin line is the 2-point correlation function averaged over 1000 realization, and the thick line is the input. The two lines overlap and it is impossible to distinguish.

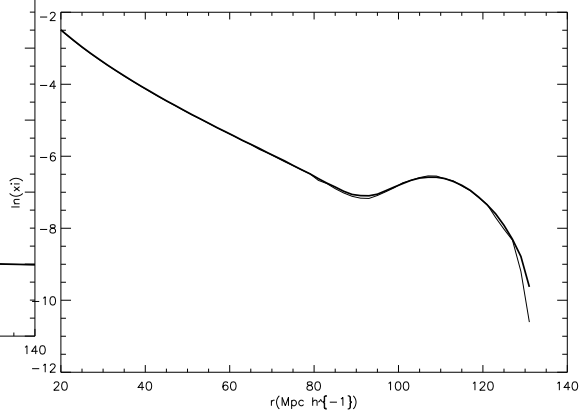

Figure 2.7 The 2-point correlation function, with a logarithmic y axis. The thin line is the 2-point correlation function averaged over 1000 realization, and the thick line is the input. We see a small discrepancy here.

2.5.3 Covariance matrix

We use the inverted covariance matrix in our calculation of χ^2(see equation (2.15)). However, the absolute value of the 2-point correlation function decreases significantly as the separation increases, and so does the difference between the input and measured 2-point correlation functions. We show how small the discrepancy gets at large separation in figure 2.6 and 2.7.

Since the absolute value of the 2-point correlation function differs by orders of magnitude at very different radii, we show in Figure 2.8 and 2.9 the correlation matrix instead of the covariance matrix. The correlation matrix is the covariance matrix

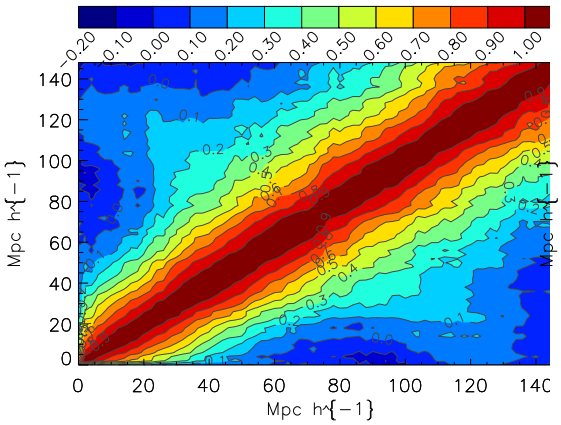

Figure 2.8 Correlation matrix of 100 realizations.

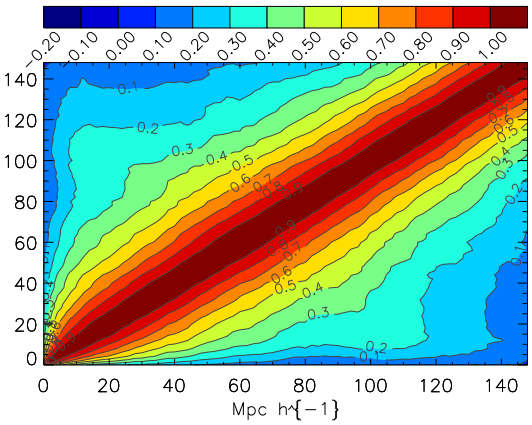

Figure 2.9 Correlation matrix of 1000 realizations.

normalized by the diagonal components of the covariance matrix, as $C_{ij}/\sqrt{C_{ii}C_{jj}}$, such that we can compare the amount of correlation with respect to its diagonal value.

We first thought we would need to use a *hybrid* covariance matrix, i.e. a covariance matrix calculated based on the power spectrum(Sánchez et al., 2008), and then normalize it by the standard deviation of the measured 2-point correlation function, yielding a smooth covariance matrix. However, by comparing figure 2.8 to 2.9, we find that the 1000 realizations provide sufficient precision for the covariance matrix that is smooth enough to calculate χ^2, so we use covariance matrix calculated directly from simulation data.

From figure 2.9, we find that different scales are correlated and there is an enhanced correlation around the scale of the BAO peak, as we expect from the shape

of the 2-point correlation function.

We should be careful about the numerical precision of the matrix inversion. We use a double-precision matrix inversion code to achieve errors on the resulting identity matrix $I = C \times C^{-1}$ down to the 10^{-15} level.

2.5.4 Uncertainty in α

For each realization, we use non-linear least-square fit code which Chi-Ting Chiang implemented to find α that gives the minimum χ^2. What the code does is as follows; From equation (2.15), we show that there are five variables that we can fit to the model. Let them be $\vec{a}$. Then, starting from a set of initial guess of parameters $\vec{a_{ini}}$, the code calculates the χ^2 at the initial guess assuming that the likelihood function follows the quadratic form around the minimum. After that, the code calculates the first derivative vector $\vec{\beta}$ and second derivative matrix $\vec{\gamma}$ of χ^2, with respect to $\vec{a}$. By multiplying $\vec{\beta}$ with the inverse of $\vec{\gamma}$, we get $\Delta\vec{a}$. By adding $\Delta\vec{a}$ to the initial guess, we get the next set of guess parameters, and we repeat the process iteratively until we get close enough to the minimum χ^2. The χ^2 calculation is done over the separation range 14 h^{-1}Mpc $< r <$ 140 h^{-1}Mpc. Figure 2.10 shows the resulting distribution of α.

From figure 2.10, we find that the dilation parameter α is within 1σ from unity. This leads us to the conclusion that our estimator is unbiased in measuring

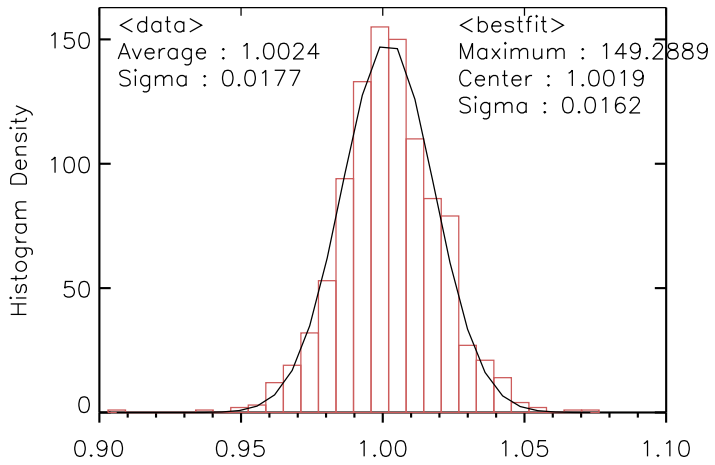

Figure 2.10 Distribution of α that gives the minimum χ^2 value for each realization. The black solid line indicates the best-fit Gaussian function. Values quoted on the upper left corner show the average and standard deviation of the histogram, and those on the upper right corner show the mean and standard deviation of the best-fit Gaussian function.

α.

2.6 2-dimensional 2-point correlation function and the Alcock-Paczyński (AP) Test

We now briefly show the 2-dimensional correlation function we test, and for the future reference, introduce how to take into account the so-called Alcock–Paczyński (AP) test (Alcock & Paczynski, 1979) using information which can be obtained from the 2-dimensional 2-point correlation function. An understanding of the red-shift space distortion is also required; however, in this thesis we assume that the

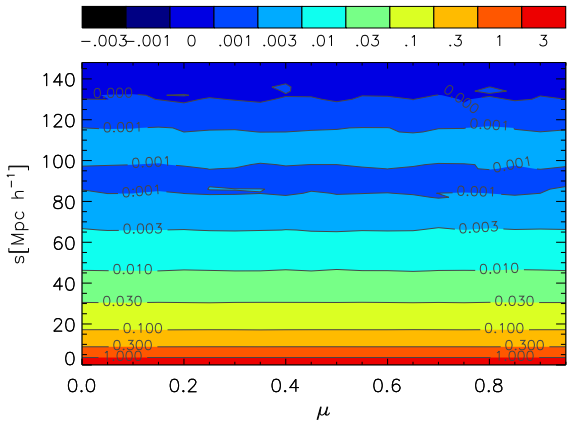

Figure 2.11 Average of 2-dimensional 2-point correlation function from 1000 realizations. The horizontal axis is cosine of the angle between the line-of-sight and the pair galaxy separation vectors, μ, and the vertical axis is the separation s.

systematics from the redshift space distortion is already taken into account, and show how we can improve the constraints on D_A and H by combining the BAO and the AP test.

First, we show the 2-dimensional 2-point correlation function in Figure 2.11. As we do not include the redshift space distortion, the 2-dimensional 2-point correlation function is expected to show no angular dependence, as we indeed find. The effects of survey geometry, selection function, as well as the redshift space distortion on the 2-dimensional 2-point correlation function is still needed to be included. Thus, as a future work, we need to study the same effects on the correlation function in two dimensions as we did in one dimension.

Then, why do we need the 2-dimensional 2-point correlation function? While

the BAO offers a robust way to determine D_A and H, the precision with which one can determine these quantities is limited because the BAO is a tiny ($\sim 2\%$) feature in the correlation function. To improve the precision, Shoji et al. (2009) have shown that using the full information contained in the power spectrum, instead of focusing on BAO only, allows us to significantly improve on the precision of D_A and H separately; this method relies on the AP test which, when combined with BAO, yields substantially more precise measurements of D_A and H, and the AP test requires information in both line-of-sight direction and the direction perpendicular to it.

In order to show this, let us remind ourselves that the 2-point correlation function is given by the volume average of density fluctuations at 2 spatially separated points

$$\xi(\vec{x}, \vec{r}) = \langle \delta(\vec{x}) \delta(\vec{x} + \vec{r}) \rangle. \tag{2.17}$$

Assuming isotropy and homogeneity of the universe, we can reduce the dependence of ξ and write it as $\xi(r)$. Due to the spherical averaging, $\xi(r)$ is sensitive only to the volume-averaged distance measure, $D_V \approx \left(\frac{D_A^2}{H}\right)^{1/3}$. The exponents of D_A and H are different because the angular separation is observed on the sky (two-dimensions) while the redshift separation is observed along the line of sight (one-dimension).

Then, how does the AP test allows us to make full use of information to separately constrain D_A and H? The AP test takes advantage of the statistical

isotropy of the universe: namely, the correlation function should only depend on the magnitude of the separation between two points in space. (Note that this statement is valid only after correcting for the redshift space distortion caused by the peculiar velocity of galaxies.) Simply put, the AP test uses the fact that length scales inferred from different directions (the line-of-sight direction and the ones perpendicular to it) must be the same.

By eliminating a length scale r_s from equations 2.3 and 2.4, one can see that the statistical isotropy demands

$$D_A(z)H(z) = \frac{c\delta z}{(1+z)\delta\theta}. \tag{2.18}$$

This equation no longer requires us to know the length scale (i.e., the standard ruler such as r_s) *a priori*. By simply demanding that the correlation function measured from the line-of-sight direction and that measured from the perpendicular direction be equal, one can determine $D_A H$ precisely.

Now that we have $D_V = \left(\frac{D_A^2}{H}\right)^{1/3}$ measured from the standard ruler (BAO) method and $D_A H$ measured from the AP test, we can combine them to determine D_A and H separately. While D_V is not as precisely determined as $D_A H$ (because it relies on a tiny BAO feature), it is precise enough to break the degeneracy between D_A and H. Figure 2.12 shows that D_V and $D_A H$ are almost orthogonal to each other, so that they are very complementary for measuring D_A and H. In this thesis,

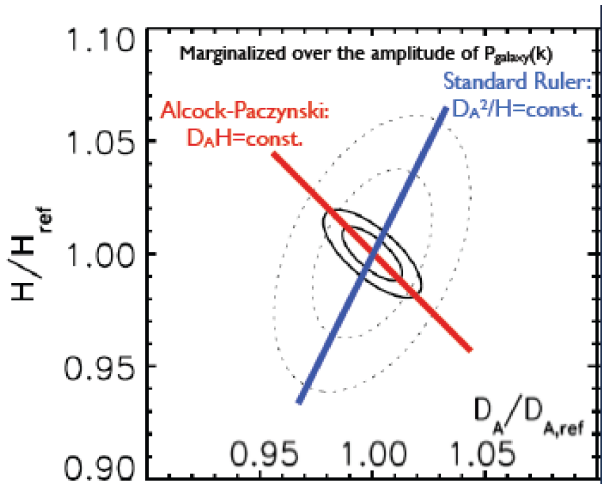

Figure 2.12 Figure 3 of Shoji et al. (2009), modified to include two lines showing the directions of the constraints on D_A and H coming from $D_V \approx D_A^2/H$ and $D_A H$. The contours show the 68% CL and 95% CL constraints expected from the HETDEX 3-year survey.

we leave the detailed analysis and the application of the AP test to constrain D_A and H as a future work.

Chapter 3

Strong Lensing as a Standard Ruler

3.1 Physical motivations

Strong gravitational lensing has long been used as a way to measure a distance-like quantity, the so-called time-delay distance $D_{\Delta t}$. The observables needed for such a measurement are the angular position of the images of the source and the time delay(which requires the source to be variable) due to a specific mass distribution. However, the time delay distance is not a real distance but rather a mathematical combination of three characteristic angular diameter distances in the source-lens system: the distance from the earth to the lens, multiplied by the distance from the earth to the source divided by the distance from the lens to the source. But consider adding one more observable, the measured velocity dispersion of the lens

galaxy. Since the time delay is determined by the gravitational potential and the velocity dispersion is determined by the mass, combining these two pieces of information yields the *actual* separation of the source light deflected by the lens galaxy. Also, from direct observation of the image positions, we can measure the *angular* separation of the deflected images. Once we know both the actual separation and the angular separation, it is straightforward to calculate the angular diameter distance $D_A(EL)$ between Earth and the lens.

3.2 Simple analysis using a singular isothermal sphere

Here we present an analysis of the simplest lens mass model that can be described as a singular isothermal sphere (SIS), which Paraficz & Hjorth (2009) considered when they first proposed measuring $D_A(EL)$ using the velocity dispersion and the time delay data.

The density distribution ρ_{SIS} for an SIS is given by

$$\rho_{SIS}(r) = \frac{\sigma^2}{2\pi G r^2}, \tag{3.1}$$

where σ is the velocity dispersion. Then, the Einstein ring radius, θ_E, is related to σ^2 via

$$\sigma^2 = \theta_E \frac{c^2}{4\pi} \frac{D_A(ES)}{D_A(LS)}, \tag{3.2}$$

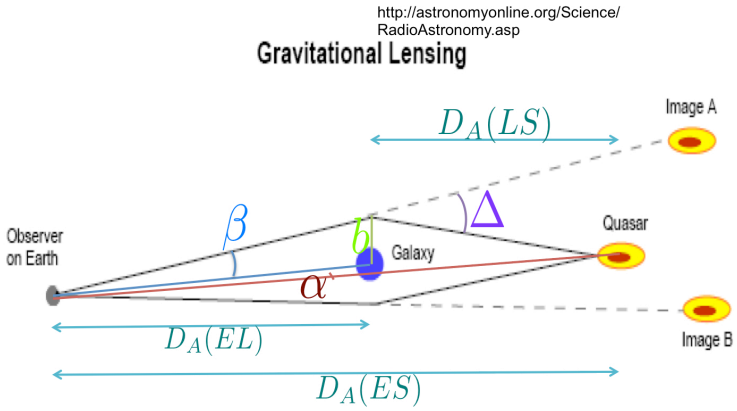

Figure 3.1 The configuration for the strong lensing system, with definition of variables used throughout the analysis. All angles are measured with respect to the center of the lens galaxy : $\vec{\beta}$ is the angular position of the image, $\vec{\alpha}$ is the angular position of the source in the absence of the lens, and Δ is the deflection angle of the light.

where $D_A(ES)$ and $D_A(LS)$ are the angular diameter distance to the source from the Earth and from the lens, respectively. See Figure 3.1 for the definition of these distances. Clearly, the relation between the two observable quantities θ_E and σ depends on the distance *ratio*.

However, we have not yet included the lensing time delay (Refsdal, 1964). The time delay is observable only when the source is temporally variable. As the light coming from the source travels through the universe to reach the observer, the presence of mass between the observer and the source, usually galaxies and clusters of galaxies, deflects the path of the light. If the mass of such lens objects is large, the bending effect will be strong enough that we observe multiple images of the one

source appearing at different positions on the sky; this phenomenon is called strong gravitational lensing.

Now, the light rays corresponding to these multiple images reach us at different times because the path length of each light ray from the source to us is different for each image. If the source is variable, we can quantify these "time delays" between images. This is called the *geometrical time delay.* Also, according to General Relativity, the existence of mass (and thus a gravitational potential) induces a gravitational time delay, which is called the *potential time delay.* One can calculate the time delay given a known mass distribution for the lens galaxy as a function of $D_A(LS)$ and $D_A(ES)$ from equation (3.2) and the angular diameter distance, $D_A(EL)$.

For a lens whose density profile is given by a singular isothermal sphere, the time delay can be written as

$$\Delta\tau = \frac{1+z_L}{2c}\frac{D_A(EL)D_A(ES)}{D_A(LS)}(\beta_2^2 - \beta_1^2) \tag{3.3}$$

where β_i is the angle between the i-th image and the lens galaxy. This equation has been used to derive time delay distance, $D_{\Delta t} \equiv (1+z_L)\frac{D_A(EL)D_A(ES)}{D_A(LS)}$, to the lens system B1608+656 (Suyu et al., 2010). However, this result depends primarily on H_0 and has a limited sensitivity to the other cosmological parameters such as the equation of state of dark energy.

Remarkably, when we combine the above equation with equation (3.2) (which are still valid only for a spherical system following a singular isothermal sphere density profile), we obtain the angular diameter distance to the lens:

$$D_A(EL)(\beta_2 - \beta_1) = \frac{c^3}{4\pi} \frac{\Delta\tau}{\sigma^2(1+z_L)}. \tag{3.4}$$

The physical interpretation of this is as follows: the time delay is determined by the potential. Combining it with the velocity dispersion, which tells us about the mass of the lens system, yields the size of the system. Since the angular scale of the system is directly observable via Einstein ring, we can use strong lensing galaxy as a standard ruler!

3.3 More realistic analysis

The above argument assumes the simplest possible lens system: a singular isothermal sphere density profile. Would the method work when considering more realistic density distributions, velocity distributions, etc?

In this section, we address this by finding more general equations for strong lensing, following the analysis of Suyu et al. (2010).

The observable corresponding to the density profile is the projected surface

density, so let us define the projected dimensionless surface density $\kappa(\vec{\theta})$ as

$$\kappa(\vec{\beta}) \equiv \frac{\Sigma(D_A(EL)\vec{\beta})}{\Sigma_{cr}}, \tag{3.5}$$

where the surface density Σ is

$$\Sigma(\vec{\beta}) = \int_{-\infty}^{\infty} \rho(\vec{\beta}, z)\ dz \tag{3.6}$$

and

$$\Sigma_{cr} \equiv \frac{c^2}{4\pi G} \frac{D_A(ES)}{D_A(EL)D_A(LS)}, \tag{3.7}$$

is the surface density within the Einstein radius.

Let the angular position of the image be $\vec{\beta}$ and that of the source be $\vec{\alpha}$. The time delay between them, $t(\vec{\beta}, \vec{\alpha})$, can be written as

$$t(\vec{\beta}, \vec{\alpha}) = \frac{1}{c}(1 + z_L)\frac{D_A(EL)D_A(ES)}{D_A(LS)}\phi(\vec{\beta}, \vec{\alpha}), \tag{3.8}$$

where ϕ is the so-called Fermat potential, defined as

$$\phi(\vec{\beta}, \vec{\alpha}) \equiv \frac{(\vec{\beta} - \vec{\alpha})^2}{2} - \psi(\vec{\beta}), \tag{3.9}$$

and ψ is the lens potential, which is calculated as

$$\psi = \frac{1}{\pi} \int d^2\beta' \kappa(\vec{\beta}) \ln|\vec{\beta} - \vec{\beta}'|. \tag{3.10}$$

We can relate this density to the lens potential via

$$\kappa(\vec{\beta}) = \frac{1}{2}\nabla^2\psi(\vec{\beta}), \tag{3.11}$$

and the lens equation becomes

$$\vec{\beta} - \vec{\alpha} = \nabla\psi(\vec{\beta}). \tag{3.12}$$

Now, defining the time-delay distance, $D_{\Delta t}$, as

$$D_{\Delta t} = (1 + z_L)\frac{D_A(EL)D_A(ES)}{D_A(LS)}, \tag{3.13}$$

then we can re-write the time-delay (equation (3.8)) as

$$t(\vec{\beta}, \vec{\alpha}) = \frac{D_{\Delta t}}{c}\phi(\vec{\beta}, \vec{\alpha}). \tag{3.14}$$

In the equation above, t is an observable, and ϕ can be modeled to satisfy observational constraints such as the image position, flux ratio and the time delay ratio between multiple pairs of images; thus, we can obtain the time delay distance. In the following analysis, we consider a power-law density profile, external convergence, and an anisotropic velocity dispersion to see which parameters influence the determination of $D_A(EL)$ the most.

3.3.1 Power-law non-isothermal models

The analysis in section 2.2 assumes a singular isothermal sphere model with an isotropic velocity dispersion. In this section, we allow the density to follow a general power-law:

$$\rho = \rho_0 (\frac{r}{r_0})^{-\gamma} \tag{3.15}$$

where γ is a free parameter. From equation (3.14), we can write the time delay between two images $\vec{\beta_i}$ and $\vec{\beta_j}$ as

$$\Delta t_{i,j} = \frac{D_{\Delta t}}{2c}(\phi(\vec{\beta_i}) - \phi(\vec{\beta_j})). \tag{3.16}$$

The lens potential can be calculated from the density equation, and a simple dimensional analysis shows that

$$\psi(\vec{\beta}) \propto |\vec{\beta}|^{-\gamma+3}, \tag{3.17}$$

such that

$$(-\gamma + 3)\psi = \vec{\beta} \cdot \nabla\psi(\vec{\beta}) = \vec{\beta} \cdot (\vec{\beta} - \vec{\alpha}). \tag{3.18}$$

By defining the slope of the lens potential to be l, i.e.

$$(-\gamma + 3) \equiv l, \tag{3.19}$$

we get

$$l\psi = \vec{\beta} \cdot (\vec{\beta} - \vec{\alpha}) \tag{3.20}$$

Now, combining the above result with equation (3.16),

$$\Delta t_{i,j} = \frac{D_{\Delta t}}{2c} \left[(\vec{\beta_i} - \vec{\alpha})^2 - (\vec{\beta_j} - \vec{\alpha})^2 - \frac{2}{l}(\vec{\beta_i} \cdot (\vec{\beta_i} - \vec{\alpha}) - \vec{\beta_j} \cdot (\vec{\beta_j} - \vec{\alpha})) \right], \tag{3.21}$$

or

$$\Delta t_{i,j} = \frac{D_{\Delta t}}{2c} \left\{ \frac{(2-l)}{l}(\beta_j^2 - \beta_i^2) + 2\frac{1-l}{l}\vec{\alpha} \cdot (\vec{\beta_i} - \vec{\beta_j}) \right\}. \tag{3.22}$$

This result is obtained by Witt et al. (2000).

Now, the mass of a lens galaxy within a radius r is given by

$$M(r) = \int_0^r 4\pi r^2 \rho_0 r_0^\gamma r^{-\gamma} dr = \frac{4\pi\rho_0 r_0^\gamma}{-\gamma + 3} r^{-\gamma+3}, \tag{3.23}$$

and producing an acceleration

$$\vec{g}(\vec{r}) = -\frac{4\pi G \rho_0 r_0^\gamma}{3-\gamma} r^{-\gamma} \vec{r} \tag{3.24}$$

on the test mass. Also, according to the post-Newtonian approximation in General Relativity, the rate of change of the direction of the velocity vector $\vec{u}$ is given as

$$c^2 \frac{d\vec{u}}{dt} = -2\vec{u} \times (\vec{u} \times \vec{g}). \tag{3.25}$$

Now let us define a new parameter Δ as the angle by which the light is deflected as it passes near the lens galaxy, and let us calculate Δ from the lens mass model we constructed previously. In the cases we consider, the deflection angle will be quite small. Thus, we can choose coordinates such that the path of the light is roughly along the x-axis, and the line connecting the center of the lens galaxy to the point of the closest approach is along the y axis. Again, because the deflection angle is small, we use the thin lens approximation, namely, light is bent suddenly at the closest approach to the lens. Thus, we can say that $r^2 = b^2 + x^2$, and, more importantly, $\vec{u} \times (\vec{u} \times \vec{r}) = c^2\vec{b}$, so that

$$\begin{aligned} \Delta &= \frac{8\pi G\rho_0 r_0^\gamma}{c^2(\gamma-3)} \int_{-\infty}^{\infty} b(r^{-\gamma})dx \\ &= \frac{8\pi G\rho_0 r_0^\gamma}{c^2(\gamma-3)} b \int_{-\infty}^{\infty} (\frac{1}{x^2+b^2})^{\gamma/2} dx \\ &= \frac{8\pi G\rho_0 r_0^\gamma b^{2-\gamma}}{c^2(\gamma-3)} \frac{\sqrt{\pi}\ \Gamma[\frac{1}{2}(-1+\gamma)]}{\Gamma(\frac{\gamma}{2})} \end{aligned} \tag{3.26}$$

where $\gamma > 1$.

From the geometry of the system we can obtain a relation among $\vec{\beta}$, $\vec{\alpha}$, and Δ as

$$\vec{\beta} - \vec{\alpha} = \frac{D_A(LS)}{D_A(ES)}\vec{\Delta} \tag{3.27}$$

where $\vec{\Delta} = \Delta \times \hat{\beta}$ since the potential only has a radial component. Then we can use

Δ in equation (3.21) to extract $D_A(EL)$ from $D_{\Delta t}$:

$$\Delta t_{i,j} = D_A(EL)\frac{(1+z_L)}{2c}\left\{(\vec{\Delta_i}+\vec{\Delta_j})(\vec{\beta_i}-\vec{\beta_j}) - \frac{2}{l}(\vec{\beta_i}\cdot\vec{\Delta_i} - \vec{\beta_j}\cdot\vec{\Delta_j})\right\}. \quad (3.28)$$

Then, how can we relate Δ to observables? Under our power-law density profile model,

$$\Delta = \frac{8\pi G\rho_0 r_0^{\gamma} b^{2-\gamma}}{c^2(\gamma-3)}F(\gamma) = -\frac{2GM(b)}{c^2 b}F(\gamma) \quad (3.29)$$

where

$$F(\gamma) \equiv \frac{\sqrt{\pi}\Gamma(\frac{1}{2}(-1+\gamma)}{\Gamma(\frac{\gamma}{2})}. \quad (3.30)$$

On the other hand, the virial theorem gives

$$\sigma_p^2 = \frac{-\gamma+3}{3(-2\gamma+5)}\frac{GM(R)}{R} \propto R^{-\gamma+2}, \quad (3.31)$$

where σ_p is the line-of-sight velocity dispersion of the system measured at R, which is $\frac{1}{\sqrt{3}}$ of the velocity dispersion σ.

Since both Δ and σ^2 are proportional to $r^{-\gamma+2}$, we can write $\Delta(b)$ as

$$\Delta = -\frac{6(5-2\gamma)}{c^2(3-\gamma)}\sigma_p^2 F(\gamma)(\frac{b}{R})^{-\gamma+2}, \quad (3.32)$$

and, from equation (3.28), $D_A(EL)$ can be obtained as

$$D_A(EL) = \frac{\Delta t_{i,j}(-\gamma+3)c^3}{(1+z_L)3\sigma_p^2 F(\gamma)(-2\gamma+5)[(\beta_i+\beta_j)(-(\frac{\beta_i}{\theta})^{-\gamma+2}+(\frac{\beta_i}{\theta})^{-\gamma+2})- \frac{2}{-\gamma+3}\{\beta_i(\frac{\beta_i}{\theta})^{-\gamma+2}-\beta_j(\frac{\beta_j}{\theta})^{-\gamma+2}\}]} \quad (3.33)$$

where θ is the angular position from the center of the lens where the velocity dispersion is measured; thus, $R = \theta D_A(EL)$.

We can also confirm that by substituting $\gamma = 2$, we can reproduce the result from the SIS model (equation (3.4)). The main difference between the SIS and the power law density profiles is that in the latter case, the velocity dispersion is a function of radii, and thus we need to take into account the radial dependence of the velocity dispersion. Note that image positions are in general, different from the point where the velocity dispersion is measured, which is why we use $\left(\frac{b}{R}\right)^{-\gamma+2}$ term.

To calculate how $D_A(EL)$ changes with γ, we take the partial derivative of

$D_A(EL)$ with respect to γ :

$$\begin{aligned}
\frac{\partial D_A(EL)}{\partial \gamma} =& \frac{\Delta t_{i,j} c^3}{3(1+z_L)\sigma_p^2}\Big[\frac{-\Gamma(0.5\gamma)}{(5-2\gamma)\sqrt{\pi}X\Gamma[0.5(-1+\gamma)]} \\
&+ \frac{2(3-\gamma)\Gamma(0.5\gamma)}{(5-2\gamma)^2\sqrt{\pi}X\Gamma[0.5(-1+\gamma)])} \\
&- \frac{(3-\gamma)\Gamma[0.5\gamma](-(\frac{2(\beta_1(\frac{\beta_1}{\theta})^{2-\gamma}-\beta_2(\frac{\beta_2}{\theta})^{2-\gamma})}{(3-\gamma)^2}}{(5-2\gamma)\sqrt{\pi}X^2\Gamma[0.5(-1+\gamma)]} \\
&+ \frac{(\beta_1+\beta_2)(-(\frac{\beta_1}{\theta})^{2-\gamma}\ln\frac{\beta_1}{\theta}+(\frac{\beta_2}{\theta})^{2-\gamma}\ln\frac{\beta_2}{\theta})))}{(5-2\gamma)\sqrt{\pi}X^2\Gamma[0.5(-1+\gamma)]} \\
&- \frac{(3-\gamma)\Gamma[0.5\gamma]\frac{2(-\beta_1(\frac{\beta_1}{\theta})^{2-\gamma}\ln\frac{\beta_1}{\theta}+\beta_2(\frac{\beta_2}{\theta})^{2-\gamma}\ln\frac{\beta_2}{\theta})}{(3-\gamma)}}{(5-2\gamma)\sqrt{\pi}X^2\Gamma[0.5(-1+\gamma)]} \\
&- \frac{(3-\gamma)\Gamma[0.5\gamma]\{\text{PolyGamma}[0,0.5(-1+\gamma)]-\text{PolyGamma}[0,0.5\gamma]\}}{(5-2\gamma)2\sqrt{\pi}X\Gamma[0.5(-1+\gamma)]}\Big] \\
=& D_A(EL)\Big[-\frac{1}{-\gamma+3}+\frac{2}{5-2\gamma}-\frac{1}{2}\int_0^\infty \frac{e^{-0.5(-1+\gamma)t}-e^{-0.5\gamma t}}{1-e^{-t}}dt \\
&- \frac{-(\frac{2(\beta_1(\frac{\beta_1}{\theta})^{2-\gamma}-\beta_2(\frac{\beta_2}{\theta})^{2-\gamma})}{(3-\gamma)^2}+(\beta_1+\beta_2)(-(\frac{\beta_1}{\theta})^{2-\gamma}\ln\frac{\beta_1}{\theta}+(\frac{\beta_2}{\theta})^{2-\gamma}\ln\frac{\beta_2}{\theta})}{X} \\
&+ \frac{\frac{2(-\beta_1(\frac{\beta_1}{\theta})^{2-\gamma}\ln\frac{\beta_1}{\theta}+\beta_2(\frac{\beta_2}{\theta})^{2-\gamma}\ln\frac{\beta_2}{\theta})}{(3-\gamma)}}{X})\Big];
\end{aligned} \tag{3.34}$$

where X is written as

$$X = (\beta_1+\beta_2)\left((\frac{\beta_1}{\theta})^{2-\gamma}-(\frac{\beta_2}{\theta})^{2-\gamma}\right) - \frac{2(\beta_1(\frac{\beta_1}{\theta})^{2-\gamma}-\beta_2(\frac{\beta_2}{\theta})^{2-\gamma})}{(3-\gamma)}. \tag{3.35}$$

When $\gamma \simeq 2$ and $\Delta\gamma \ll 1$, the above equation becomes

$$\begin{aligned}\frac{\Delta(D_A(EL))}{D_A(EL)} \simeq \Big[& -\frac{0.2695}{\beta_1-\beta_2} \\ & -\frac{0.07958\Big\{-2(\beta_1-\beta_2)+(\beta_1+\beta_2)\Big(-\ln\big(\frac{\beta_1}{\theta}\big)+\ln\big(\frac{\beta_2}{\theta}\big)\Big)}{(\beta_1-\beta_2)^2} \\ & +\frac{2\Big(-\beta_1\ln\big(\frac{\beta_1}{\theta}\big)+\beta_2\ln\big(\frac{\beta_2}{\theta}\big)\Big)\Big\}}{(\beta_1-\beta_2)^2}\Big]\Delta\gamma.\end{aligned} \tag{3.36}$$

This is still for a spherically symmetric model of the lens galaxy. In the next section we perturb the model by using an anisotropic velocity dispersion, parametrized by an anisotropic radius r_{ani}, which is related to the ratio of tangential to radial velocity dispersions.

3.3.2 Anisotropic velocity dispersion

The next step is to include an anisotropic velocity dispersion parameter on top of the spherical stellar distribution. Suyu et al. (2010) use simple spherical Jeans modeling to relate the observable line-of-sight velocity dispersion to the mass distribution, assuming that the stellar distribution follows the Hernquist profile. Here, the anisotropy parameter r_{ani} is given by

$$\beta_{ani} = \frac{r^2}{r_{ani}^2 + r^2} = 1 - \frac{\sigma_T^2}{\sigma_r^2}, \tag{3.37}$$

where σ_T and σ_r are the velocity dispersions in the tangential and radial direction, respectively. The anisotropy is parametrized by only one variable r_{ani} in the above

solution, but by linearly superimposing the solutions, we are able to model almost any velocity structures (Merritt, 1985). This is related to the luminosity-weighted projected velocity dispersion, σ_s, by

$$I_H(R)\sigma_s^2 = 2\int_R^\infty \Big(1-\beta_{ani}\frac{R^2}{r^2}\Big)\frac{\rho_*\sigma_r^2 r dr}{\sqrt{r^2-R^2}} \tag{3.38}$$

where the three-dimensional velocity dispersion, σ_r, is given by

$$\begin{aligned}\sigma_r = \frac{4\pi G a^{-\gamma}\rho_0}{3-\gamma}\frac{r(r+a)^3}{r^2+r_{ani}^2}\Big(&\frac{r_{ani}^2}{a^2}\frac{{}_2F_1[2+\gamma,\gamma;3+\gamma;\frac{1}{1+r/a}]}{(2+\gamma)(r/a+1)^{2+\gamma}}\\ &+\frac{{}_2F_1[3,\gamma;1+\gamma;-a/r]}{\gamma(r/a)^\gamma}\Big).\end{aligned} \tag{3.39}$$

The Hernquist profile ρ_* (Hernquist, 1990) is given by

$$\rho_*(r) = \frac{I_0 a}{2\pi r(r+a)^3}, \tag{3.40}$$

and the projected Hernquist distribution, $I_H(R)$, is known to provide a good fit for the stellar distribution of elliptical galaxies that follow the de Vaucouleurs law

$$I_H(R) = \frac{I_0}{2\pi a^2(1-s^2)^2}[(2+s^2)X(s)-3]. \tag{3.41}$$

Here, $s = R/a$ is scaled projected radius where R is the projected radius, and $X(s)$ is defined as

$$X(s) = \begin{cases} \dfrac{1}{\sqrt{1-s^2}} \operatorname{sech}^{-1} s & \text{for } 0 \le s \le 1 \\ \dfrac{1}{\sqrt{1-s^2}} \sec^{-1} s & \text{for } 1 \le s < \infty \end{cases} . \tag{3.42}$$

Using these formulae, we can rewrite equation (3.38) as

$$\begin{aligned} \frac{I_0}{2\pi a^2 (1-s^2)^2}[(2+s^2)X(s) - 3]\sigma_s^2 =& \frac{I_0 a}{2\pi} \frac{4\pi G a^{-\gamma} \rho_0}{3-\gamma} \\ & 2\int_R^\infty dr \Big(1 - \frac{R^2}{r_{ani}^2 + r^2}\Big) \frac{r}{(r_{ani}^2 + r^2)\sqrt{r^2 - R^2}} \\ & \Big(\frac{r_{ani}^2}{a^2} \frac{{}_2F_1[2+\gamma, \gamma; 3+\gamma; \frac{1}{1+r/a}]}{(2+\gamma)(r/a+1)^{2+\gamma}} \\ & + \frac{{}_2F_1[3, \gamma; 1+\gamma; -a/r]}{\gamma (r/a)^\gamma} \Big), \end{aligned} \tag{3.43}$$

which gives

$$\sigma_s^2 = \frac{8\pi G a^{-\gamma+2} \rho_0 (1-s^2)^2}{(3-\gamma)[(2+s^2)X(s) - 3]} I = \frac{GM(a)}{a} \frac{(1-s^2)^2}{(2+s^2)X(s) - 3} I \tag{3.44}$$

where

$$\begin{aligned} I \equiv & \int_R^\infty dr \Big(1 - \frac{R^2}{r_{ani}^2 + r^2}\Big) \frac{ar}{(r_{ani}^2 + r^2)\sqrt{r^2 - R^2}} \\ & \Big(\frac{r_{ani}^2}{a^2} \frac{{}_2F_1[2+\gamma, \gamma; 3+\gamma; \frac{1}{1+r/a}]}{(2+\gamma)(r/a+1)^{2+\gamma}} + \frac{{}_2F_1[3, \gamma; 1+\gamma; -a/r]}{\gamma (r/a)^\gamma} \Big). \end{aligned} \tag{3.45}$$

Here, I cannot be calculated analytically, so we evaluate it numerically with parameters for the strong lens system B1608+656 (Suyu et al., 2010). With $r_{\rm eff} = 0.58''$

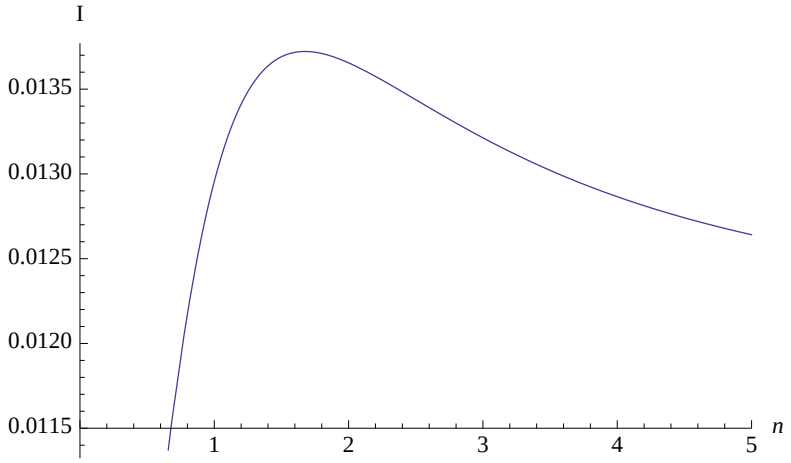

Figure 3.2 Values of I as a function of anisotropy radii, $r_{\rm ani} = nr_{\rm eff}$.

and $a = 0.551 r_{\rm eff}$, and parameterizing $r_{\rm ani} = nr_{\rm eff}$ (n is ranging from 0.5 to 5), we calculate I as a function of n : See Fig. 3.2.

Also, σ_p^2 can be calculated as

$$\sigma_p^2 = \frac{\int_A I_H \sigma_s^2 R \; dR \; d\theta}{\int_A I_H R \; dR \; d\theta} = \frac{I_0 \frac{GM(a)}{a^3} * 2 \int_0^{r_{\rm eff}} I(R) \; R \; dR}{(1/3) I_0}. \tag{3.46}$$

As $r_{\rm ani}$ dependence only comes from I, we find

$$\frac{\partial \sigma_p^2}{\partial r_{\rm ani}} = \frac{GM(a)}{a^3} B \frac{\partial I}{\partial r_{\rm ani}} = \sigma_p^2 (\frac{a}{r_{\rm eff}})^{-\gamma+2} \frac{B}{a^2} \frac{\partial I}{\partial r_{\rm ani}}, \tag{3.47}$$

where B is a numerical factor given by the integral, $B \equiv \int_0^{r_{eff}} 2I(R) \; R \; dR$, and we show $\frac{\partial I}{\partial n}$ as a function of n in Fig. 3.4.

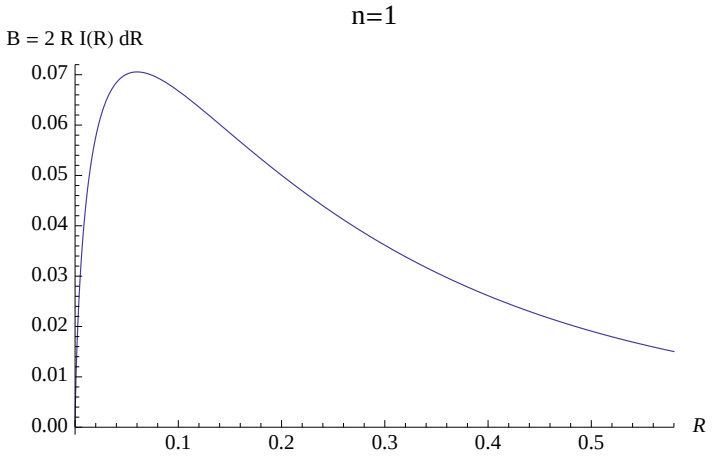

Figure 3.3 Value of B depending on R

3.3.3 External convergence

In modeling realistic lens systems, one important factor to consider is the so-called "mass-sheet degeneracy", which is an intrinsic degeneracy of any lens mass models. Once we find a model for the surface density distribution $\kappa_{model}(\vec{\beta})$ that matches the observations, we can still fit most of the observables (e.g. image positions, flux ratio for point sources, image shape for extended sources) with transformed surface density distribution, $\kappa_{trans}(\vec{\beta})$, given by

$$\kappa_{trans}(\vec{\beta}) = \lambda + (1 - \lambda)\kappa_{model}(\vec{\beta}). \tag{3.48}$$

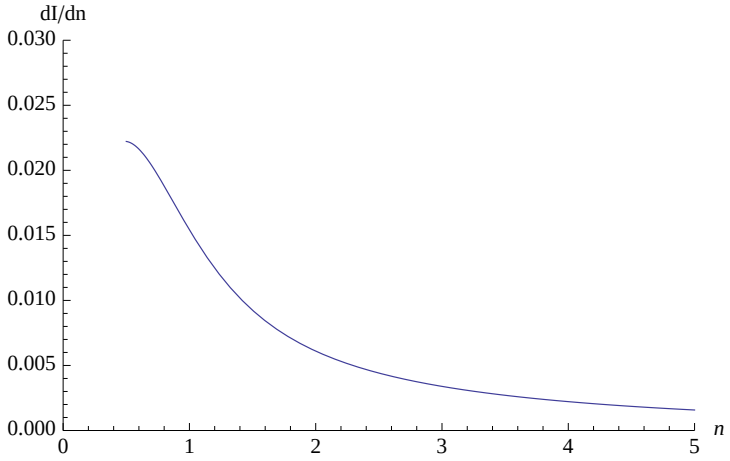

Figure 3.4 Values of $\frac{\partial I}{\partial n}$ as a function of anisotropy radii, $r_{ani} = nr_{eff}$.

This is shown by Falco et al. (1985), where the parameter λ physically corresponds to the convergence of the sheet. Considering the relation among κ , ϕ and ψ,

$$\phi = [\frac{1}{2}(\vec{\beta} - \vec{\alpha})^2 - \psi], \tag{3.49}$$

$$\nabla^2\psi = 2\kappa, \tag{3.50}$$

we find that the difference in the Fermat potential between the two images i and j also transforms as

$$\Delta\phi_{trans,i,j} = (1 - \lambda)\Delta\phi_{model,i,j}. \tag{3.51}$$

If we only consider the effect of the external convergence, then we can substitute λ in the above equation with κ_{ext}, which is defined as

$$\kappa_{ext} = \frac{4\pi G}{c^2} \sum_i \frac{\Sigma_i(D_i\vec{\theta})D_i D_{is}}{D_s}. \tag{3.52}$$

Combining the above equation with equation (3.21) yields

$$\begin{aligned}\Delta t_{i,j} =& \frac{D_{\Delta t}}{2c}(1-\kappa_{ext})\Big[\frac{(\gamma-1)}{(3-\gamma)}(\beta_j^2-\beta_i^2) + 2\frac{\gamma-2}{-\gamma+3}\vec{\alpha}\cdot(\vec{\beta_i}-\vec{\beta_j})\Big] \\ =& (1-\kappa_{ext})D_A(EL)\frac{(1+z_L)}{2c}\left\{(\vec{\Delta_i}+\vec{\Delta_j})(\vec{\beta_i}-\vec{\beta_j}) - \frac{2}{l}(\vec{\beta_i}\cdot\vec{\Delta_i}-\vec{\beta_j}\cdot\vec{\Delta_j})\right\}.\end{aligned} \tag{3.53}$$

Thus, we find that $\Delta t_{i,j}$ is simply increased by a factor of $1-\kappa_{ext}$ due to the external shear, where Δ indicates the total deflection angle. However, when we use the transformed quantities in equation (3.33) to calculate $D_A(EL)$, the $1-\kappa_{ext}$ dependence cancels out. More precisely, the density normalization, ρ_0, transforms as

$$\rho_{0,trans} = (1-\kappa_{ext})\rho_{0,model}, \tag{3.54}$$

and thus among the total deflection angle Δ, only a $(1-\kappa_{ext})$ fraction of it is from the lens. That is to say,

$$(1-\kappa_{ext})\Delta = \frac{6(-2\gamma+5)}{c^2(-\gamma+3)}\sigma_p^2 F(\gamma)(\frac{b}{R})^{-\gamma+2}, \tag{3.55}$$

and thus $D_A(EL)$ can still be simply calculated as equation (3.33). A physical explanation is as follows. All the lensing observables are affected by not only the mass inside the lens galaxy, but also the mass lying along the line-of-sight, which is the external convergence that causes the mass-sheet degeneracy. However, the dynamics of the lens galaxy has nothing to do with the mass outside the galaxy: It is solely determined by the mass inside the effective radius. Thus, by measuring the velocity dispersion of the galaxy, which is not relevant to the external convergence, we can calculate the mass inside the galaxy exclusively. Another information on mass which is given by deflection angle depends on the external convergence. Thus, by combining these two pieces of information on mass, we can break the mass-sheet degeneracy.

This is a good news, as in previous studies of the time-delay distance to measure the Hubble constant, κ_{ext} prevents us from measuring H_0 without knowing κ_{ext}(Suyu et al., 2010). On the other hand, we have shown that the angular diameter distance from strong lensing, which combines the time-delay, image positions, and velocity dispersion, does not suffer from the external convergence degeneracy.

3.4 Error formula and implications for B1608+686 and RJX1131-1231

The author's adviser, Eiichiro Komatsu, asked Sherry Suyu to apply this angular diameter distance measurement to a lens system B1608+656 from which she had measured the time-delay distance (and hence H_0) (Suyu et al., 2010). To his surprise, the inferred D_A to the lens was very uncertain, but it was not, at that point, clear which observables needed to be improved in order to make this method competitive. The goal of this part of the second-year project is to find the conditions under which this method works well. Suyu's algorithm is numerical and highly complex, which makes it difficult to easily know which observables are limiting the precision of the inferred D_A. This motivates our studying this problem for more analytical (but simpler, though still sufficiently realistic) systems.

So far, we have calculated the effects of γ, κ_{ext}, and r_{ani} on D_A. These quantities are related to the observables such as $\Delta t_{i,j}$, σ_p^2, and $\beta_{i,j}$ in a complicated way. Assuming that these observables are independently measured, we write the total uncertainty on $D_A(EL)$, hereafter S_{D_A}, as

$$S_{D_A} = \sqrt{(\frac{\partial D_A}{\partial \Delta t_{i,j}})^2 S^2_{\Delta t} + (\frac{\partial D_A}{\partial \sigma_p^2})^2 S^2_{\sigma_p^2} + (\frac{\partial D_A}{\partial \gamma})^2 S^2_{\gamma} + (\frac{\partial D_A}{\partial r_{ani}})^2 S^2_{r_{ani}}} \tag{3.56}$$

$$= D_A(EL)\sqrt{(\frac{1}{\Delta t_{i,j}})^2 S^2_{\Delta t} + (\frac{1}{\sigma_p^2})^2 S^2_{\sigma_p^2} + \frac{1}{D_A^2}(\frac{\partial D_A}{\partial \gamma})^2 S^2_{\gamma} + \frac{1}{D_A^2}(\frac{\partial D_A}{\partial r_{ani}})^2 S^2_{r_{ani}}} \tag{3.57}$$

Image positions, $\beta_{i,j}$, are precisely measured, thus we do not include the uncertainty of it in this formula. While κ_{ext} does not appear in the estimated value of D_A, the uncertainty in D_A depends on κ_{ext}. We find

$$S_{D_A} = D_A(EL)\left[(1-\kappa_{ext})^2\{(\frac{1}{\Delta t_{i,j}})^2 S^2_{\Delta t} + (\frac{1}{\sigma_p^2})^2 S^2_{\sigma_p^2} + \frac{1}{D_A^2}(\frac{\partial D_A}{\partial r_{ani}})^2 S^2_{r_{ani}}\} + \frac{1}{D_A^2}(\frac{\partial D_A}{\partial \gamma})^2 S^2_{\gamma}\right]^{1/2}. \tag{3.58}$$

where $(\frac{\partial D_A}{\partial r_{ani}})$ should be evaluated as

$$\frac{\partial D_A}{\partial r_{ani}} = \frac{\partial D_A}{\partial \sigma_p^2}\frac{\partial \sigma_p^2}{\partial r_{ani}} \tag{3.59}$$

which gives

$$S_{D_A} = D_A(EL)\left[(1-\kappa_{ext})^2\{(\frac{1}{\Delta t_{i,j}})^2 S^2_{\Delta t} + (\frac{1}{\sigma_p^2})^2\left(S^2_{\sigma_p^2} + (\frac{\partial \sigma_p^2}{\partial r_{ani}})^2 S^2_{r_{ani}}\right)\} + \frac{1}{D_A^2}(\frac{\partial D_A}{\partial \gamma})^2 S^2_{\gamma}\right]^{1/2}. \tag{3.60}$$

We calculate derivatives, $\frac{\partial \sigma}{\partial r_{ani}}$ and $\frac{\partial D_A}{\partial \gamma}$, numerically. In the following sections, we apply this formula to two lens systems, B1608+656 and RXJ1131-1231.

3.4.1 B1608+656

Figure 3.5 shows the image configuration of B1608+656 (Suyu et al., 2009). The information on image configuration is important as our formula applies only to a circularly symmetric case. Thus, the only image pairs we can use are the ones that

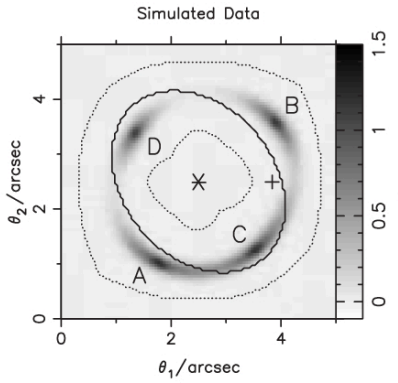

Figure 3.5 Image of B1608+656, adopted from Figure 1 of Suyu et al. (2009). The solid line and the dotted line are not relevant to our study.

are on the opposite sides of the lens center. The data of B1608+656 are summarized

as :

$$
\begin{aligned}
z_L =& 0.6304 \\
z_s =& 1.394 \\
\beta_A =& 1.6502'' \\
\beta_B =& 1.7743'' \\
\beta_C =& 1.6184'' \\
\beta_D =& 1.4560'' \\
r_{eff} =& 0.58'' \\
a =& 0.551 r_{eff} \\
\gamma =& 2.08 \pm 0.03 \\
\sigma_p =& 260 \pm 15 \ kms^{-1} \\
\Delta t_{AB} =& 31.5^{+2.0}_{-1.0} \ days \\
\Delta t_{CB} =& 36.0^{+1.5}_{-1.5} \ days \\
\Delta t_{DB} =& 77.0^{+2.0}_{-1.0} \ days \\
\Delta t_{CD} =& \Delta t_{CB} - \Delta t_{DB} = -41.0^{+2.5}_{-1.8} \ days.
\end{aligned}
\tag{3.61}
$$

Using these values, we find $D_A(EL) = 1391.58$ Mpc, and $D_{\Delta t} = 5424.85$ Mpc. For comparison, D_A and $D_{\Delta t}$ from the best-fit WMAP7 parameters are $D_A(EL) =$

1423.1 Mpc and $D_{\Delta t} = 5093.86$ Mpc. $D_{\Delta t}$ is calculated from equation (3.53),

$$D_{\Delta t} = \frac{2c\ \Delta t_{i,j}}{(1-\kappa_{ext})\left[\frac{\gamma-1}{3-\gamma}(\beta_j^2-\beta_i^2) + 2\frac{\gamma-2}{-\gamma+3}\vec{\alpha}\cdot(\vec{\beta_i}-\vec{\beta_j})\right]}. \tag{3.62}$$

Suyu et al. (2009) estimated the external convergence as:

$$\kappa_{ext} = 0.10^{+0.08}_{-0.05}. \tag{3.63}$$

Note that the uncertainty of the time delay is different for each of the image pairs, but the uncertainty caused by the external convergence is the same for two pairs. Thus, we first need to average the uncertainties of the time delay measurements, and then include the effect of the external convergence. Calculation of the two pairs of S_D, only considering the time delay, yields :

$$D_{\Delta t,AB} = 5290.0^{+335.87}_{-167.93}\ \text{Mpc}, \tag{3.64}$$

$$D_{\Delta t,CD} = 5530.5^{+337.19}_{-242.78}\ \text{Mpc}. \tag{3.65}$$

Now we can calculate inverse variance weighted average of $D_{\Delta t}$, and add uncertainties from the external convergence to obtain

$$\hat{D_{\Delta t}} = 5424.85^{+554.45}_{-347.41}\ \text{Mpc}. \tag{3.66}$$

Our estimator, which is based on a simpler mass model, is able to reproduce the result of Suyu et al. (2009) reasonably well. We can convert this measurement into H_0. Our estimate is

$$\frac{\Delta H_0}{H_0} = \frac{\Delta D_{\Delta t}}{D_{\Delta t}} \simeq^{+9.37}_{-5.87} \%, \tag{3.67}$$

while Suyu et al. (2010) find $D_{\Delta t} = 5160^{+290}_{-240}$ Mpc and $H_0 = 70.6^{+3.1}_{-3.1} km\ s^{-1}$ Mpc^{-1}, which is $\approx 4.39\%$. The two results are in good agreement with each other considering the simplicity of our model, and with the same data.

Now we can numerically assess the uncertainty on $D_A(EL)$, using equation (3.60). As n varies in the range $0.5 < n < 5$, we can define the uncertainty in n as $\delta n = 4.5$, thus $S_{r_{ani}} = \delta n \times r_{eff} = 2.61''$. Also, from equation (3.34), $\frac{1}{D_A(EL)} \frac{\partial D_A}{\partial \gamma} = 2.083$. Thus,

$$\begin{aligned} S_{D_A} = \ & D_A \Bigg\{ 3.905 \times 10^{-3} + (1-\kappa_{ext})^2 \times \Big[(\frac{1}{\Delta t_{i,j}})^2 S^2_{\Delta t} + \\ & (\frac{1}{\sigma_p^2})^2 \left(S^2_{\sigma_p^2} + (0.551^{-\gamma+2} \sigma_p^2 \frac{B}{a^2} \frac{\partial I}{\partial (n r_{eff})})^2 (S_{r_{ani}})^2 \right) \Big] \Bigg\}^{1/2} \\ = \ & D_A \Bigg\{ 3.905 \times 10^{-3} + (1-\kappa_{ext})^2 \times \Big[4.031 \times 10^{-3} + \\ & (\frac{1}{\sigma_p^2})^2 \left((0.1154 \times \sigma_p^2)^2 + (0.551^{-\gamma+2} \sigma_p^2 \frac{B}{a^2} \frac{\partial I}{\partial n})^2 (2.61)^2 \right) \Big] \Bigg\}^{1/2} \\ \simeq \ & D_A \sqrt{3.905 \times 10^{-3} + 3.265 \times 10^{-3} + 1.079 \times 10^{-2} + 5.491 \times 10^{-5}} \end{aligned} \tag{3.68}$$

where from the first term, each number indicates the fractional uncertainty on D_A contributed by the density profile index γ, the time delay measurement $\Delta t_{i,j}$, the line-of-sight velocity dispersion measurement σ_p, and the anisotropic radius $r_{\rm ani}$.

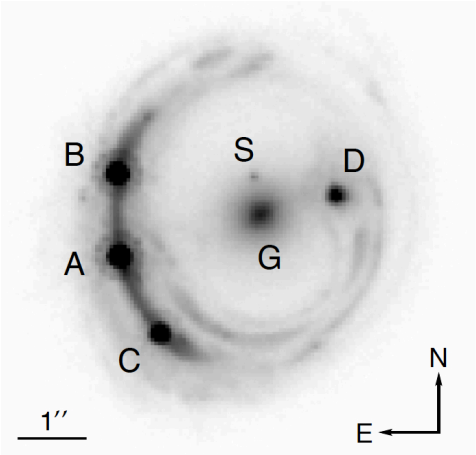

Figure 3.6 Image of RXJ1131-1231, adopted from Figure 1 of Suyu (2012).

Thus, from our numerical analysis, and assuming $r_{\rm ani} = nr_{eff}$, we find that the prior on the velocity anisotropy, $0.5 < n < 5$, makes small contribution to the uncertainty on D_A. The most dominant source of uncertainty is the velocity dispersion measurement. The other contributions are also comparable. Among these contributions, contribution by γ depends largely on the lens system configuration, as we show next.

3.4.2 RXJ1131-1231

In this section we repeat the same analysis as above, but with another well-studied strong lens system, RXJ1131-1231. The data for this system are

$$
\begin{aligned}
z_L &= 0.295 \\
z_s &= 0.658 \\
\beta_A &= 2.102'' \\
\beta_D &= 1.131'' \\
r_{eff} &= 1.85 \\
\gamma &= 1.95^{+0.05}_{-0.04} \\
\sigma_p &= 323 \pm 20 \ km/s \\
\Delta t_{AB} &= 0.7 \pm 1.4 \ days \\
\Delta t_{DB} &= 91.4 \pm 1.5 \ days \\
\Delta t_{AD} &= \Delta t_{AB} - \Delta t_{DB} = -90.7 \pm 2.1 \\
\kappa_{ext} &= 0.089^{+0.006}_{-0.006}.
\end{aligned}
\tag{3.69}
$$

Using these values, we find $D_A(EL) = 911.0$ Mpc, and $D_{\Delta t} = 2450.8$ Mpc. For comparison, D_A and $D_{\Delta t}$ from the best-fit WMAP7 parameters are $D_A(EL) = 876.5$ Mpc and $D_{\Delta t} = 2506.86$ Mpc. We show the equivalent of figures 3.3 and 3.4 for this system in figure 3.7 and 3.8. Repeating the previous analysis, we find

$$
S_{D_A} \simeq \ D_A \sqrt{3.409 * 10^{-4} + 4.445 * 10^{-4} + 1.273 * 10^{-2} + 6.927 * 10^{-6}}. \tag{3.70}
$$

For RXJ1131-1231, the dominant source of uncertainty is again the velocity

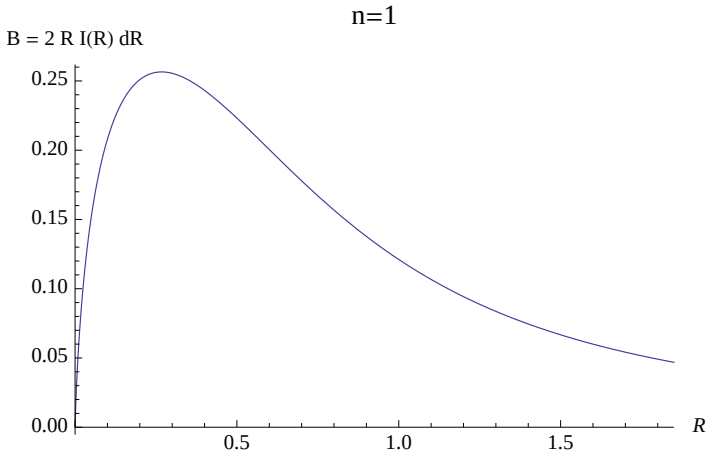

Figure 3.7 Values of B as a function of R, for RXJ1131-1231.

dispersion. On the other hand, the contributions from the uncertainties in the time delay and γ are smaller than those for B1608+656 by an order of magnitude. This is due to the difference in image configurations. The lens galaxy of B1608+656 is located near the center of the Einstein ring, i.e., $\beta_1 - \beta_2 \ll \beta_2$. On the other hand, the lens galaxy of RXJ1131-1231 is displaced from the center of Einstein ring, giving $\beta_1 - \beta_2 \simeq \beta_2$. Now, as it follows equation (3.34) that the derivative of D_A with respect to γ is inversely proportional to $\beta_1 - \beta_2$ and $(\beta_1 - \beta_2)^2$, the contribution from the uncertainty in γ is significantly amplified for B1608+656. Also, while the uncertainties in the time-delay measurements are similar for both lens systems, the central values of the time delay are bigger for RXJ1131-1231 by a factor of three, as images A and B of RXJ1131-1231 are located at different radii from the lens galaxy, and thus they receive more different delays.

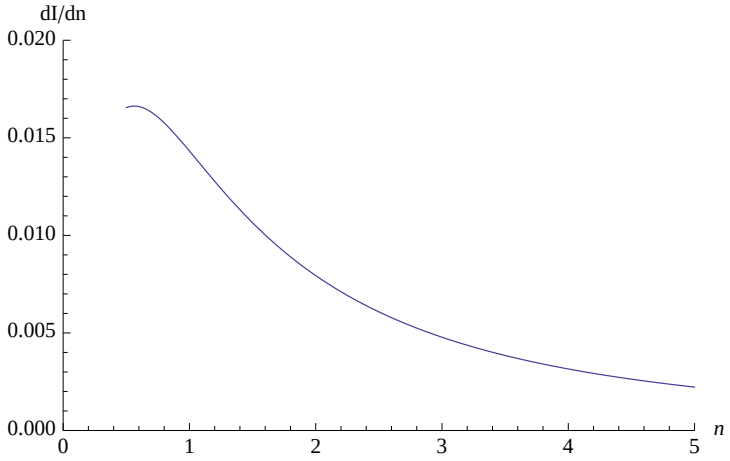

Figure 3.8 Values of $\frac{\partial I}{\partial n}$ as a function of anisotropy radius, $r_{ani} = nr_{eff}$, for RXJ1131-1231.

Figure 3.9 shows the posterior distribution of D_A inferred from RXJ1131-1231 as a function of the uncertainties in the velocity dispersion.[1] Here, we artificially increase or reduce the uncertainties in the velocity dispersion, to see how they affect the posterior distribution of D_A. This analysis goes beyond our simple analysis presented in this thesis, and is based upon the analysis pipeline used by Suyu et al. (2012). We find that the uncertainty in the inferred value of D_A decreases as the uncertainty in the velocity dispersion decreases, although the rate at which the uncertainty in D_A decreases is slower than expected. For example, a factor of two reduction in the uncertainty in the velocity dispersion gives only 10% reduction in the uncertainty in D_A. At the moment it is not clear as to why the improvement in D_A is slower than we expect. Understanding this requires further study.

[1] Sherry Suyu, private communication.

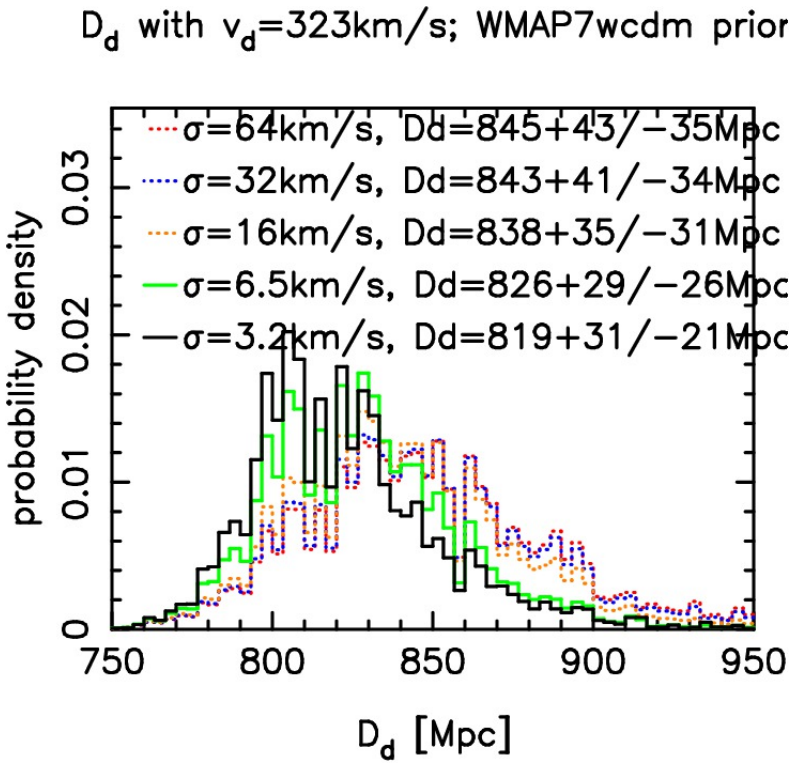

Figure 3.9 Posterior probability density distribution of the angular diameter distance to the lens galaxy of RXJ1131-1231, as a function of the uncertainty in the velocity dispersion. The measured central value of the velocity dispersion is 323 km/s, and the lines show the uncertainties of 20%, 10%, 5%, 2%, and 1%, respectively. The values of the angular diameter distance (quoted as D_d in the figure) show the medians of credible intervals based on the 16th and 84 percentiles.

3.5 Conclusion

In this thesis we have studied two completely different ways to measure the angular diameter distances.

First, we have investigated how accurately the HETDEX survey can determine the angular diameter distance using the BAO feature in the 2-point correlation

function of LAEs. Using realistic simulations incorporating the HETDEX survey geometry and mask, we find that the 2-point correlation function is not biased by the survey geometry and mask unlike the power spectrum, and the 2-point correlation function of the HETDEX data in a low-redshift bin ($1.9 < z < 2.5$) and the northern sky field can determine the angular diameter distance with 2% accuracy.

Second, we have investigated how we can use strongly lensed systems to determine the angular diameter distances. We have extended a simple analysis done by Paraficz & Hjorth, who use a spherical singular isothermal sphere, by including spherical mass distribution with an arbitrary power-law slope; an anisotropic velocity dispersion; and the external convergence. We derive the angular diameter distance as a function of various observables such as the time-delay, image positions, power-law index of the less mass distribution, the velocity dispersion, and an anisotropic velocity dispersion parameter. We find that the external convergence cancels out, and thus it does not affect the determination of the angular diameter distance. By differentiating this formula with respect to the observables, we derive the formula for the uncertainty in D_A as a function of the uncertainties in the observables. Applying this formula to the strong lens systems B1608+656 and RXJ1131-1231, we conclude that the uncertainty in D_A inferred from these systems should be dominated by the uncertainty in the velocity dispersion.

We would like to acknowledge Eiichiro Komatsu, Donghui Jeong, Chi-Ting

Chiang, and Edward L. Robinson for their advises, comments and discussions on academic aspects of this work, as well as help on resolving practical problems of it. Funding for this work has been provided in parts by the University of Texas College of Natural Sciences and Texas Cosmology Center through grants to Eiichiro Komatsu and Karl Gebhardt, and in part by the National Science Foundation award AST-0926815 for Hobby-Eberly Telescope Dark Energy Experiment.

Bibliography

Alcock, C., & Paczynski, B. 1979, Nature, 281, 358

Anderson, L., et al. 2012, MNRAS, 427, 3435

Chiang, C.-T., Wullstein, P., Jeong, D., Komatsu, E., Blanc, G. A., et al. 2013

Coe, D., & Moustakas, L. A. 2009, ApJ, 706, 45

Dobke, B. M., King, L. J., Fassnacht, C. D., & Auger, M. W. 2009, MNRAS, 397, 311

Eisenstein, D. J., et al. 2005, ApJ, 633, 560

Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1

Fukugita, M., Futamase, T., & Kasai, M. 1990, MNRAS, 246, 24P

Fukugita, M., & Turner, E. L. 1991, MNRAS, 253, 99

Futamase, T., & Yoshida, S. 2001, Progress of Theoretical Physics, 105, 887

Hernquist, L. 1990, ApJ, 356, 359

Hill, G. J., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Jeong, D., & Komatsu, E. 2006, ApJ, 651, 619

—. 2009, ApJ, 691, 569

Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64

Merritt, D. 1985, AJ, 90, 1027

Oguri, M. 2007, ApJ, 660, 1

Ohyama, Y., et al. 2002, AJ, 123, 2903

Paraficz, D., & Hjorth, J. 2009, A&A, 507, L49

Refsdal, S. 1964, MNRAS, 128, 307

Sánchez, A. G., Baugh, C. M., & Angulo, R. E. 2008, MNRAS, 390, 1470

Shoji, M., Jeong, D., & Komatsu, E. 2009, ApJ, 693, 1404

Suyu, S. H. 2012, MNRAS, 426, 868

Suyu, S. H., Marshall, P. J., Auger, M. W., Hilbert, S., Blandford, R. D., Koopmans, L. V. E., Fassnacht, C. D., & Treu, T. 2010, ApJ, 711, 201

Suyu, S. H., Marshall, P. J., Blandford, R. D., Fassnacht, C. D., Koopmans, L. V. E., McKean, J. P., & Treu, T. 2009, ApJ, 691, 277

Witt, H. J., Mao, S., & Keeton, C. R. 2000, ApJ, 544, 98

Yamamoto, K., & Futamase, T. 2001, Progress of Theoretical Physics, 105, 707

Yamamoto, K., Kadoya, Y., Murata, T., & Futamase, T. 2001, Progress of Theoretical Physics, 106, 917